

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

AN ALTERNATIVE STATISTICAL MEASURE FOR RACIALLY POLARIZED VOTING

Kenneth McCue and Arthur Lupia



SOCIAL SCIENCE WORKING PAPER 690

January 1989

Abstract

Measurements of the existence and extent of racially polarized voting are often at the forefront of the evidence presented in vote dilution litigation. Previous models used in the measurement of racially polarized voting have been inadequate for a broad range of cases. The source of this inadequacy is that these models were not based upon assumptions about individual behavior.

A new model, which is based on reasonable assumptions about individual behavior and can be approximated by a varying parameters model, is derived. By contrasting the new model with the other models, it is shown that both the correlation coefficient and linear regression can lead to inaccurate, misleading or incorrect conclusions about the state of racial polarization in the electorate. After providing a straightforward statistical test for identifying misspecification and showing how to obtain estimates for this model by the method of maximum likelihood, the model is compared with an individual level data set of an election where racially polarized voting occurred (1988 California presidential primary). A new measure for the estimation of racially polarized voting is then proposed.

One of the most significant changes in the amended Voting Rights Act of 1982 affected the type of evidence minority groups could use when attempting to overturn a vote-diluting electoral mechanism. Since then, the relevant legal standard of any action under the Act has been the “results test”.¹ Under this standard, minority groups can support their vote dilution claims by proving the existence of racially polarized voting. Consequently, measurements of the existence and extent of racially polarized voting are often at the forefront of the evidence presented in vote dilution litigation.²

It is our contention that the methods currently used to measure racially polarized voting can be inadequate, inaccurate, misleading or incorrect. These problems arise directly because these methods are utilized without any knowledge of how they relate to the actual voting behavior of individuals in the electorate. In particular, the estimation of voting behavior from aggregate data requires more than the Pearson correlation coefficient and more than a multiple regression whose dependent variables describe only the size of minority groups. Unless ethnic identity is the only determinant of voting behavior, both of these methods can lead to statistical inferences which are inconsistent with the actual voting behavior of members of minority groups.

Our goal in this article, then, is to provide the theoretical basis for developing a method of measuring racially polarized voting. This theoretical basis consists of assumptions on the behavior of the individual in the electorate, from which we utilize basic probability theory to derive a statistically estimable model. In the process, we utilize this model to examine previous measures of racially polarized voting, and provide insight into how they relate to the voting behavior of individuals in the electorate.

Specifically, we develop in this paper:

1. The necessary conditions for estimation.

¹Senate Report # 97-417, 97th Congress, Second Session 28 (1982), U.S. Code Cong. and Admin. News 1982, pp. 177-205.

²Among the most recent and/or prominent: *Major v. Treen*, 574 F. Supp. 325 (1983), *Thornberg v. Gingles*, 106 S. Ct. 2752 (1986), *Smith v. Clinton*, U.S. District Court (Arkansas), Civil Action LRC 88-29

2. How to estimate racially polarized voting with this model.
3. How this model can be derived from basic probabilistic reasoning on the nature of analyzing aggregate vote returns.
4. How this model can be approximated by a type of model known in the statistical literature as the *varying parameters* model.
5. The exact correspondence between this model and the correlation coefficient in the determination of racially polarized voting.
6. What happens when the necessary conditions are violated in linear regression.
7. How we can use properties of the model to assess both its consistency and the underlying assumptions that our model's use requires.
8. How to estimate this model.
9. A comparison of this model with an individual level data set on an election where racially polarized voting takes place.
10. A new measure of racially polarized voting.

1 Necessary conditions for estimation.

Consider an electorate consisting of $g = 1, \dots, G$ groups satisfying the partition condition.³

Partition condition:

The groups must be defined so that they are mutually exclusive and collectively exhaustive. (Each individual in one and only one group.)

³We appreciate the advice of Bruce Cain who has suggested the utility of partitioning the electorate in communications with the authors.

A “group” could be one of the various ethnic or racial groups in the electorate, or a subdivision of such a group based on geographic and/or socioeconomic differences within an ethnic classification. Furthermore, let there be $c = 1, \dots, C$ candidates running in a particular election for which members of this electorate vote. Let there be $u = 1, \dots, U$ aggregated units for which members of each group g vote. For example, these “units” can be thought of as any electoral subdivision, like voting precincts. Let the number of voters of group g in aggregate unit u be X_{ug} . Then, if V is the total number of votes cast in this election, and V_u is the total number of votes cast in each electoral subdivision, we have,

$$V = \sum_{u=1}^{u=U} V_u \text{ where } V_u = X_{u1} + \dots + X_{uG}.$$

The model we present is based on the following simple condition:

Homogeneity condition:

Every individual member of group $g = 1, \dots, G$ votes with a probability p_{gc} for candidate c . This probability is independent of other voters’ decisions and electoral unit.

This condition states that when an individual is selected from a group, the probability of him voting for a candidate is the same as the probability of any other individual chosen from that group voting for the candidate. Thus an individual’s behavior can be described, probabilistically, as a Bernoulli trial. The assumption of independence is no different than that made in survey analysis.⁴

The *homogeneity* condition must be satisfied for any model of group voting behavior, using correlation or regression techniques for estimation, to always be correct. A model not satisfying this condition is liable to serious misspecification. This misspecification will be troubling, in the context of vote dilution

⁴It is difficult to see how any method of the analysis of aggregate voting, which is based on a description of individual behavior, can be accepted without the utilization of some form of the homogeneity condition, for what it does is assign a probability law to the individual himself. Gary King has suggested in a communication with the authors that this assumption be relaxed somewhat by allowing the probability to be randomly distributed as well. For reasons which will become apparent in section 3 and 4 this is possible but if it is done much of the simplicity of our approach will be lost. We discuss this more in the conclusion.

litigation, since analysis with the misspecified model can lead to one of these two results, (Detailed in Section 6):

1. Existence of racially polarized voting established when, in fact, it does not exist.
2. Non-existence of racially polarized voting established, when, in fact it does exist.

2 How to estimate racially polarized voting with this model

It may seem that partitioning the electorate into different groups would make it impossible to determine the voting behavior of a broad group which consists of two or more partitioned subgroups, but this is not the case. In general, it is true that there will not be a “single” probability for a group as broad as “blacks” or “hispanics”. However, there is a straight-forward method of obtaining a single probability for this type of group. The method is the same as that used in public opinion polls. First, calculate the sum of the obtained probabilities for each subgroup, weighting each probability by the number of voters within the partition. Then normalize this figure, dividing the weighted sum by the total number of voters in all of the partitioned groups. This gives us the single probability for the broad group, which is needed for questions of racial polarization.

We want a model which satisfies the *homogeneity* and *partition* conditions. There always exists a finite set of divisions of the electorate which satisfy the partition condition, for at an extreme we can create groups with one person in each, which would give mutual exclusivity and collective exhaustion.⁵ When these conditions are satisfied we proceed to the estimation of the voting behavior of

⁵Of course, this much reduction would not permit the model to be estimated. We might mention at this point that while verification of the partition condition is relatively straightforward, verification of the homogeneity condition is more difficult. In Section 7, we provide a statistical test which will provide evidence on whether the conditions of the model are satisfied.

a particular ethnic group.

We can always divide an ethnic group, E , into relevant mutually exclusive and collectively exhaustive subgroups in order to analyze voting behavior.⁶ After verifying the satisfaction of the partition and homogeneity conditions, we estimate the probability that an *individual* in each subgroup will vote for candidate c ($c = 1, \dots, C$). We desire an estimate of the probability that a randomly selected member of ethnic group E will vote for candidate c . To do this we sum the individual probabilities in the manner described above, which gives the expected vote for candidate c from ethnic group E .

If there are G groups in the electorate, where groups 1 through ℓ are mutually exclusive, collectively exhaustive subgroups of ethnic group E , then we have

$$V_c = \frac{\text{Subgroups of } E}{X_1 p_{c1} + X_2 p_{c2} + \dots + X_\ell p_{c\ell} + X_{\ell+1} p_{c\ell+1} + \dots + X_G p_{cG}}.$$

Our estimate for the probability that a member of ethnic group E votes for candidate c is thus:

$$p_{Ec} = \frac{\sum_{g=1}^{\ell} \sum_u X_{ug} p_{cg}}{\sum_{g=1}^{\ell} \sum_u X_{ug}}.$$

If the probabilities are random (as will be the case when the model is statistically estimated), we can use a straight-forward application of probability theory to determine the standard error of this probability.

3 The basic probabilistic framework

Using the *homogeneity* condition, it is a simple exercise in probability theory to note that, by the binomial theorem, if we have X_{ug} individuals that are members of group g in election unit u , the probability of K of them voting for candidate c , $\Pr(K_{ugc} = k)$, is simply,

⁶For example, within ethnic group E there are differences in income, education, age, etc., all of which may have some effect on the (heterogeneity) homogeneity of the voting behavior of the ethnic group.

$$\Pr(K_{ugc} = k) = \binom{X_{ug}}{k} p_{cg}^k (1 - p_{cg})^{X_{ug} - k}. \quad (1)$$

Thus, the probability that candidate c will receive V_{uc} votes from election unit u , $\Pr(V_{uc} = K_{u1c} + \dots + K_{uGc})$ is

$$\begin{aligned} & \sum_{\ell_1=0}^{V_{uc}} \Pr(K_{u1c} = \ell_1) \left[\sum_{\ell_2=0}^{V_{uc}-\ell_1} \Pr(K_{u2c} = \ell_2) \dots \left[\sum_{\ell_{G-1}=0}^{V_{uc}-\sum_{k=1}^{G-2} \ell_k} \Pr(K_{u(G-1)c} = \ell_{G-1}) \right. \right. \\ & \quad \left. \left. \Pr(K_{uGc} = X_{uG} - \sum_{m=1}^{G-1} \ell_m) \underbrace{\llbracket \dots \rrbracket}_{G-2 \text{ times}} \right] \right] \end{aligned}$$

which equals

$$\begin{aligned} & \sum_{\ell_1=0}^{V_{u1}} \binom{X_{u1}}{\ell_1} p_{1c}^{\ell_1} (1 - p_{1c})^{V_{u1} - \ell_1} \left[\sum_{\ell_2=0}^{V_{uc}-\ell_1} \binom{X_{u2}}{\ell_2} p_{2c}^{\ell_2} (1 - p_{2c})^{V_{uc}-\ell_2} \dots \right. \\ & \quad \left[\sum_{\ell_{G-1}=0}^{V_{uc}-\sum_{k=1}^{G-2} \ell_k} \binom{X_{u(G-1)}}{\ell_{G-1}} p_{(G-1)c}^{\ell_{G-1}} (1 - p_{(G-1)c})^{V_{uc}-\sum_{k=1}^{G-2} \ell_k} \dots \right. \\ & \quad \left. \binom{X_{uG}}{X_{uG} - \sum_{m=1}^{G-1} \ell_m} p_{Gc}^{X_{uG} - \sum_{m=1}^{G-1} \ell_m} (1 - p_{Gc})^{\sum_{m=1}^{G-1} \ell_m} \underbrace{\llbracket \dots \rrbracket}_{G-2 \text{ times}} \right] \quad (2) \end{aligned}$$

(By convention, if $\ell_i > X_{ui}$ we set the probability equal to zero.)

This is a formidable expression but there is a relatively good way of approximating it, as we show in the next section.

4 An approximation with the varying parameters model

It is shown in elementary textbooks in probability ⁷ that the cumulative distribution function of the binomial distribution may be approximated by that of

⁷Feller, William, *An Introduction to Probability Theory and its Applications: Volume I*, (New York, 1950)

the normal c.d.f.. This suggests:

$$K_{ugc} \sim n \left(X_{ug}p_{gc}, X_{ug}p_{gc}(1-p_{gc}) \right). \quad (3)$$

Define ϵ to be a normal variable,⁸

$$\epsilon_{ugc} \sim n \left(0, p_{gc}(1-p_{gc}) \right) = n(0, \lambda_{gc}). \quad (4)$$

Then we express K_{ugc} as,

$$K_{ugc} = X_{ug}p_{gc} + X_{ug}^{\frac{1}{2}} \epsilon_{ugc}. \quad (5)$$

Implying directly that the probability of $(V_{uc} = K_{u1c} + \dots + K_{uGc})$ can be approximated by,

$$\Pr(V_{uc} = X_{u1}p_{1c} + X_{u1}^{1/2}\epsilon_{u1c} + \dots + X_{uG}p_{Gc} + X_{uG}^{1/2}\epsilon_{uGc}), \quad (6)$$

or, collecting terms,

$$\Pr(V_{uc} = X_u p_c + Y_u \epsilon_{uc}), \quad (7)$$

where X , p and ϵ are now vectors and Y_u is the vector consisting of the $X_{ug}^{1/2}, g = 1, \dots, G$. This is a variant of a *varying parameters* model.⁹ Note that this formulation shows that linear regression is a consistent (though not efficient) method of estimation.

Since

$$\Pr(V_{uc} - X_u p_c = Y_u \epsilon_{uc}), \quad (8)$$

we have,

$$\Pr(Y_u \epsilon_{uc}) = \frac{1}{(2\pi)^{1/2} (X_u \Lambda_c)^{1/2}} \exp \left[-\frac{1}{2} \frac{(Y_u \epsilon_{uc})^2}{X_u \Lambda_c} \right] \quad (9)$$

⁸This approximation is considered to be excellent when $X_{ug}p_{gc} \geq 5$ and $X_{ug}(1-p_{gc}) \geq 5$. It is often quite good for smaller values. It is clear, by the construction of the error structure in this model, that the error terms are independent of each other and the variables X_{ug}

⁹See Gregory C. Chow, "Random and Changing Coefficient Models" in *Handbook of Econometrics: Volume 2*, (Amsterdam, 1984), pp. 1214 - 1224. and P. A. V. B. Swamy, *Statistical Inference in Random Coefficient Regression Models*, (New York, 1971)

$$\Pr(V_{uc} - X_u p_c) = \frac{1}{(2\pi)^{1/2} (X_u \Lambda_c)^{1/2}} \exp \left[-\frac{1}{2} \frac{(V_{uc} - X_u p_c)^2}{X_u \Lambda_c} \right],$$

where $\Lambda_c = [\lambda_{1c}, \dots, \lambda_{Gc}]$.

The product of these probabilities for the individual observations gives us the likelihood. We then utilize the method of maximum likelihood to obtain estimates of the unknown parameters $p_{1c}, \dots, p_{Gc}, \lambda_{1c}, \dots, \lambda_{Gc}$.

5 The correspondence between this model and the correlation coefficient.

We now look at the correspondence between this model and the correlation coefficient in the case where there exist two groups that satisfy the homogeneity and partition conditions. The correlation coefficient is a measure of the relationship between the proportion of the total vote recieved by a candidate c , V_c , and the proportion of the population who are members of a particular group, X_g (one such group could be “hispanic” voters). The true equation for V_c , as the sum of the expected vote percentage of all groups in the electorate, is:

$$V_c = X_1 p_{1c} + X_2 p_{2c}$$

One problem with the correlation coefficient is that its value depends not only on the propensity of a particular group to vote for a particular candidate but also on factors which may have little or nothing to do with the issue of voting polarization. Given the true formulation of the vote for candidate c above, we derive in Appendix A that the correlation coefficient is actually,

$$\text{Corr}(V_c, X_1) = \frac{p_{1c} - p_{2c}}{[(p_{1c} - p_{2c})^2 + a_1 p_{1c}(1 - p_{1c}) + a_2 p_{2c}(1 - p_{2c})]^{1/2}},$$

$$\text{where } a_i = \frac{E(X_i^2)}{\text{Var}(X_i)}.$$

Thus, while the correlation coefficient is increasing as the difference between p_{1c} and p_{2c} increases ¹⁰, the actual relation between $\text{Corr}(V_c, X_1)$ and p_{1c} is non-linear and dependent upon the moments of X_1 and X_2 . Stated less elegantly, when p_{1c} or p_{2c} changes, the resulting value of $\text{Corr}(V_c, X_1)$ will not, in general, reflect the magnitude of the change. Furthermore, the magnitude of the correlation coefficient will not accurately reflect the magnitude of the difference between p_{1c} and p_{2c} .

In Appendix B, we give an example of how misleading the correlation coefficient can be. We look at the case where $p_{1c} = .5$ and $p_{2c} = 0$. We show that $\text{Corr}(X_1, V_1)$ can take values of between 0 and .65, depending on the variance of X_1 . Thus, it is possible to have a correlation coefficient which tells of “significant racial polarization” when in fact *a majority of neither group supported the candidate in question*. In cases where the homogeneity and partition conditions are not satisfied, interpreting the correlation coefficient will be even more problematic.

Grofman, Migalski and Noviello (1985) make the following comment about the correlation coefficient:

“there has been no universally accepted threshold at which polarization would be regarded as statistically or substantively significant”

but point out that the courts have accepted the standard that $\rho \geq .5$ implies that “*statistically significant racial polarization*” exists. ¹¹ The preceding argument should make clear the danger of such interpretations of the correlation coefficient. We assert that this section adds to previous findings which conclude that the correlation coefficient should no longer be considered as a reasonable method of estimation for voting rights litigation.¹²

¹⁰It requires some algebra, but differentiate the expression with respect to p_1 . It can be shown that for any fixed value of p_{2c} , as p_{1c} increases so does $\text{Corr}(V_c, X_1)$.

¹¹Grofman, Bernard, Michael Migalski and Nicholas Noviello, “The Totality of Circumstances” test in Section 2 of the 1982 Extension of the Voting Rights Act: A Social Science Perspective”, *Law and Policy*, Vol. 1, Number 2 (1985), pp. 199-223. Quote is on page 203.

¹²Engstrom, Richard L. and Michael D. MacDonald “Quantitative Evidence in Vote Dilution Litigation, Part II: Minority Coalitions and Multivariate Analyses”, *Urban Lawyer*, Winter 1987, Vol. 19 # 1, p. 65 -75. It is shown in this article that the correlation coefficient should not be used when “more than two minority

6 Violation of assumptions in linear regression.

We now directly assess the consequences of using linear regression to estimate racially polarized voting when the homogeneity and/or partition conditions are violated. We begin by constructing two simplified electorates. Each electorate consists of 100 voters (alternatively, the number of voters can be thought of as the percent of voters in an electorate of size N), and is split into four subdivisions (precincts) of 25 voters each.

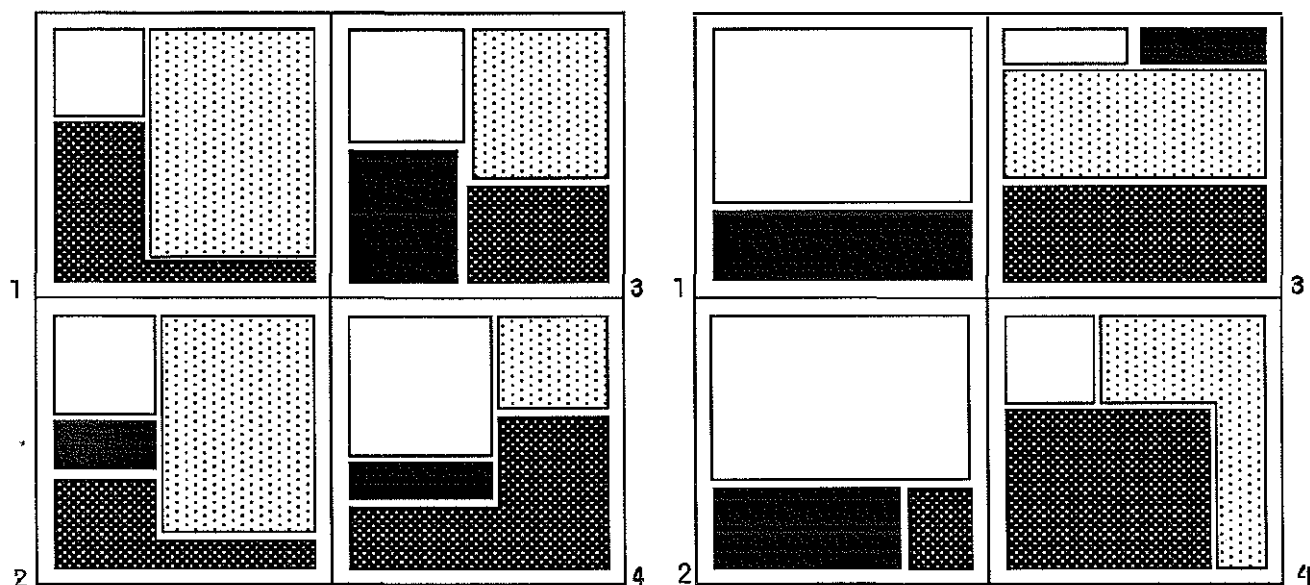
Consider Electorates U and V, detailed in Figure 1, and suppose that the only two factors which affect the voting decision of an individual are E , a racial (or ethnic) classification, and I , a socioeconomic (or geographic) classification. Furthermore, suppose that individuals can easily be placed into one of two categories within each classification. [$E = (E_1, E_2)$ and $I = (I_1, I_2)$].¹³ Thus, each voter can be described by membership in one (and only one) of these four groups: [$E_1I_1, E_1I_2, E_2I_1, E_2I_2$]. Geographic segregation by factors E and I exists in both electorates to different degrees.

The electorates are divided into four electoral units of equal size. The proportion of each group within electoral units 1 - 4, (z_1, z_2, z_3, z_4), varies across electoral units within each electorate. The shadings within the electoral units are proportional to the number of each group within that electoral unit. The distributions (among units) of the four groups in the four electoral units are represented graphically and numerically in Figure 1.

p_{gc} gives the probability that a member of group g will vote for candidate c . By looking at the table in Figure 1, you can verify that, in Electorate U, p_{gc} is

groups" are present as the resulting coefficients may be "systematically biased", depending on the covariance of the dependent variables. What we show here extends this result to the case of two groups, even if the groups satisfy our conditions. In addition, Grofman, Migalski and Noviello have advocated the use of ecological regression techniques in a series of articles including "Effects of Multimember Districts on Black Representation in State Legislatures", *The Review of Black Political Economy*, (1986), Vol. 14, Number 4, pp. 66-78. and the article referenced in the previous footnote. Among published work surveying estimation methods for racial polarization are Charles S. Bullock, III and Susan A. MacManus, "Measuring Racial Bloc Voting is Difficult for Small Jurisdictions", *National Civic Review*, July/ August, 1984

¹³These suppositions are made only to simplify the example. They are not conditions on which the general results depend.



Electorate U

Electorate V

Group Pattern	Group Name	Electorate U		Electorate V	
		Pct of Electorate	$p_{(gc)}$	Pct of Electorate	$p_{(gc)}$
	$E_1 1$	20	.9	40	.3
	$E_1 2$	40	.3	20	.2
	$E_2 1$	10	.9	15	1.0
	$E_2 2$	30	.3	25	.6

$p_{(gc)}$ is the probability of support for candidate c by a member of group g .

Group Distributions within electoral units.								
Group Name	Electorate U				Electorate V			
	$z1$	$z2$	$z3$	$z4$	$z1$	$z2$	$z3$	$z4$
$E_1 1$.12	.16	.20	.32	.72	.68	.08	.12
$E_1 2$.60	.52	.32	.16	.00	.00	.44	.36
$E_2 1$.00	.08	.24	.08	.28	.24	.08	.00
$E_2 2$.28	.24	.24	.44	.00	.08	.40	.52

Figure 1 : Description of the sample electorates.

determined entirely by factor I. In Electorate V, both factors are determinants of p_{gc} .

In Table 1, we compare three estimates of the effect of group membership on voting behavior. The first model is an exit poll, which asks each person not only who they voted for but also with which group are they a member. If proper sampling techniques are used, we directly obtain the desired relationship. The second and third models are linear regression models which have already been proposed as measures of racially polarized voting, but which violate either the partition or homogeneity conditions.¹⁴

Table 1 - VARIOUS MODELS

SAMPLE ELECTORATE	Electorate U	Electorate V
Model 1 (Exit Poll)		
E_1	.50	.27
E_2	.45	.75
I_1	.90	.49
I_2	.30	.37
Estimation from Model 2: $V_1 = \beta_1 E_1 + \beta_2 E_2$ (Violates Homogeneity)¹⁵		
E_1	.19	.59
E_2	.91	.26
Estimation from Model 3: $V_1 = \alpha_1 E_2 + \alpha_2 I_2$ (Violates Partition)		
E_2	.00	.16
I_2	-.60	-.20
Constant	22.5	12.1

¹⁴In electorates more complex than U and V, the problems we expose will be many orders of magnitude greater.

¹⁵The constant term is omitted from the estimation of Model 2 as the dependent variables are mutually exclusive and collectively exhaustive. A constant term is included in the estimation of Model 3 as the right hand side variables do not account for every individual in the electorate. Variances and covariances between the groups are provided at the beginning of the appendix.

A comparison of the estimates from Models 2 and 3 with the results of the exit poll (Model 1) exemplify the type of problem arises when the homogeneity and partition conditions are violated.

Let us first look at Model 2, $V_1 = \beta_1 E_1 + \beta_2 E_2 + u$, which is suggested in Engstrom and McDonald (1987).¹⁶ This line of reasoning also applies to the ecological regression method suggested in Grofman, Migalski and Noviello (1985) as it is possible that neither satisfies the homogeneity condition.¹⁷

According to the exit poll, 50 % of E_1 and 45 % of E_2 support candidate c in Electorate U. The coefficients from the regression using Model 2 gives evidence of substantial racial polarization in Electorate U, $(\beta_1 = .19, \beta_2 = .91)$, the wrong result given that E has no effect on p_{gc} , the individual's voting decision in that electorate. In Electorate V, 27 % of E_1 and 75 % of E_2 actually support candidate c. The use of Model 2 leads to the conclusion that E_1 is much more supportive of candidate c than E_2 , again clearly the wrong result. When using Model 2, we omit the effect of the factor I on the vote V_c , which is the cause of severe misspecification. This misspecification occurs because it is not the case that every member of ethnic group E_1 (or E_2) votes for candidate c with equal probability (i.e., $p_{(E_1 I_1 c)}$ is not equal to $p_{(E_1 I_2 c)}$).

Model 3, $V_c = \alpha_1 E_2 + \alpha_2 I_2$, suggested by the U.S. in *Thornberg v. Gingles*, may not satisfy the partition condition.¹⁸ This model makes it difficult, if not impossible, to model the behavior of individuals within a group since one

¹⁶Their model is a simple linear regression. Notice that, intuitively, we can make the homogeneity assumption operational in this model by setting the probability of voting for candidate c in group $E_1 I_1$ equal to the probability of voting for c in group $E_1 I_2$, resulting in both probabilities equalling β_1 . A more formal operationalization of this model will follow.

¹⁷In the example on pp.204-205, the partition condition is satisfied for blacks and whites, as the electorate by race and "race voted for", thus capturing the underlying electoral behavior of each group. Whether or not this model also satisfies the homogeneity condition depends on whether the behavior of the minority is homogeneous. Grofman et. al., incidentally, recognize the possibility of misspecification in their estimations and suggest looking at voting units which are racially homogenous to avoid the possibility of the "ecological fallacy", which is the rather unfortunate term attached to misspecification in aggregate data analysis. This does not preclude, of course, the possibility of different behavior among members of a racial minority in a homogeneous voting unit (as opposed to members of a non-homogeneous voting unit).

¹⁸In the context of the examples above. Changing the subscripts of E and I give the other representations of this model.

person can easily fit into multiple categories.¹⁹ This occurrence will lead to a constant term and coefficients whose interpretations are not obvious, as can be gathered from the results above.

A problem which both Model 2 and Model 3 share is specification error. A brief explanation will expose the problem. The true equation, which gives us estimates of racial voting polarization is

$$V_c = X_1 p_{1c} + X_2 p_{2c} + X_3 p_{3c} + \cdots + X_G p_{Gc}$$

where $g = 1, \dots, G$ are the partitions of the electorate. In regression analysis, p_{gc} is represented by β_{gc} . We estimate the vector of β_{gc} , β by $\hat{\beta}$, where $V_c = X\beta$ and

$$E(\hat{\beta}) = (X'X)^{-1}X'V_c = (X'X)^{-1}X'X\beta = \beta$$

By violating the partition or homogeneity conditions, the actual partitions (the X_i 's) are replaced by another set of variables, (Z_i 's). To estimate the relation between V_c and Z , we must, necessarily replace the vector of coefficients β by a vector τ , which we estimate by $\hat{\tau}$. τ can be statistically sound, but it is no longer representative of the effect of race on the vote for a particular candidate. Consider the expected value of τ :

$$E(\hat{\tau}) = (Z'Z)^{-1}Z'V_c = (Z'Z)^{-1}Z'X\beta$$

Recall that β is the true value of the vector of the p'_{gc} s. The expected value of $\hat{\tau}$ does not equal β . Furthermore, any covariance between Z and X that is not accounted for in the formulation of the Z 's will cause τ to be different than β , in ways which we may not be able to account for.

Given that the homogeneity and partition conditions are satisfied, and the type of problem just discussed does not occur, we proceed with a further examination of the situation presented in Engstrom and MacDonald (1987). They are interested in the case where there are multiple ethnic groups and propose an ordinary least squares method of regression to estimate coefficients for each of the ethnic groups (blacks, latinos and whites). Their dependent variable, "number of votes for candidate 1" will be represented as V_1 . Their dependent variables

¹⁹For instance, one person can be Hispanic, and middle-class and have a high school diploma simultaneously.

are L , number of latinos and B , number of blacks. The “number of whites” variable is not used as a dependent variable, the coefficient for this group is given by the constant term in the equation. The equation they propose is as follows:

$$V_1 = \alpha_0 + \alpha_1 L + \alpha_2 B.$$

Where α_0 , the constant term, is actually, $\alpha_0 = \alpha_0(L + B + W)$ (expressing things in terms of proportions, thus the sum of all the groups in the electorate is one). This gives:

$$V_1 = \alpha_0(L + B + W) + \alpha_1 L + \alpha_2 B$$

$$V_1 = \alpha_0 W + (\alpha_1 + \alpha_0)L + (\alpha_2 + \alpha_0)B.$$

If dividing the electorate into blacks, latinos and whites satisfies the partition condition, the true equation for this situation looks like this:

$$V_1 = p_0 W + p_1 L + p_2 B.$$

Thus we get the following equalities,

$$\alpha_0 = p_0,$$

$$\alpha_1 + \alpha_0 = p_1,$$

$$\alpha_2 + \alpha_0 = p_2,$$

$$\alpha_1 = p_1 - \alpha_0,$$

$$\alpha_2 = p_2 - \alpha_0,$$

$$\alpha_1 = p_1 - p_0,$$

$$\alpha_2 = p_2 - p_0.$$

Thus, under the regression method proposed by Engstrom and MacDonald, what is being estimated is not the voting behavior of a particular ethnic group but the voting behavior of one group relative to that of another. ²⁰

A benefit of the model proposed in this paper is that its basis is the true equation and the coefficients, β_0 , β_1 and β_2 are obtained directly from the estimation. Thus our coefficients now give us the magnitude of each group’s voting polarization, from which we can easily derive the relative magnitudes given in the Engstrom and MacDonald coefficients. ²¹ It is fairly straightforward to

²⁰A similar argument is derived in Grofman, Migalski and Noviello (1985).

²¹Consider an example where $\alpha_1 = .3$. We have shown that this represents the difference in the voting behaviors

extend this model to cases of three or more racial groups as long as the subgroups whose coefficients are estimated satisfy the homogeneity and partition conditions.

7 Assessing the Consistency of the Model

One benefit of this model is that it provides a straightforward statistical test for identifying misspecification. The binomial distribution is a one-parameter member of the exponential family, so once we know the location parameter value (call it p), we also know the variance (which is $p(1-p)$). From maximum likelihood theory, we know that minus the inverse of second partial derivatives with respect to the parameters (i.e. the information matrix) will give us the asymptotic variance. We can then construct a Wald test for a misspecified model.²² In examples of racially polarized voting, this can be caused by the exclusion of factors which do affect the voting decision, but are not distributed randomly across race, and usually lead to a violation of the homogeneity condition (i. e. income, education, age).²³ The only remedy for such misspecification that we can offer at the moment, within the context of our model, is to further subdivide the electorate into mutually exclusive and collectively exhaustive groups which account for these “excluded factors” as well as ethnicity.

For construction of the Wald test, we simply take

$$\chi^2_{(G)} \sim \left(\begin{array}{c} \hat{\lambda}_1 - p_1(1 - p_1) \\ \vdots \\ \hat{\lambda}_G - p_G(1 - p_G) \end{array} \right)' \left[E\left[\frac{-\delta^2}{\delta\Theta^2}\right]^{-1} \right] \left(\begin{array}{c} \hat{\lambda}_1 - p_1(1 - p_1) \\ \vdots \\ \hat{\lambda}_G - p_G(1 - p_G) \end{array} \right),$$

where the matrix in the middle is the information matrix for the variance pa-

of groups 0 (whites) and group 1 (latinos). In vote dilution cases, it is important to know whether this value represents the difference between .0 and .3 or the difference between .35 and .65.

²²Wald, A. “Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large”, *Transactions of the American Mathematical Society* Vol. 54, pp. 426-482.

²³It can also be caused by an incorrect assignment of individuals into their correct partition, but that is a problem which can be investigate through non-statistical means. If this occurs, however, it should cause a high Wald and a rejection of the model, also.

rameters.²⁴

The value from this test is used then to determine the outcome of a hypothesis test where the critical value is determined by a chi - squared distribution with degrees of freedom G . It is also possible to derive a likelihood ratio test which will be asymptotically equivalent to the Wald test. In that case the restricted maximization would set each estimated variance parameter equal to the product of the estimated probability parameter and one minus that parameter, and the usual likelihood ratio test would be performed.

8 How to estimate this model.

To obtain estimates for this model by the method of maximum likelihood, it is necessary to maximize the value of the likelihood function (derived in Section 4) over the parameter space. The parameter space for the model is defined as follows: the probability parameters for the individual groups fall between zero and one and the variance parameters are greater than zero. Maximum likelihood estimation can take place either utilizing a search model or a modified Newton-Raphson algorithm.

One search algorithm which is convenient for most social science researchers is contained in SPSSX. The constrained nonlinear regression procedure can be utilized to perform the maximization of a likelihood function, including the likelihood for this model. An example of this is provided in Appendix C. Our experience with this is that it utilizes a tremendous amount of computer time, but in these days of declining cpu costs this is probably not an overriding consideration.

The likelihood can also be maximized by using a Newton-Raphson procedure. To do this, we calculate the matrix of the second partials of the log of the

²⁴It is also possible to construct Wald tests for the probability parameters and the overall model, using the appropriate submatrices of the information matrix, and we utilize those later on in our empirical testing section.

likelihood, ²⁵ and obtain a matrix where the $(j, \ell)^{th}$ element in the matrix is:

$$(j, \ell) = \begin{cases} -\sum_{u=1}^U \frac{X_{uj}X_{u\ell}}{\sigma_u^2} \left[\frac{1}{2} - \frac{(V_{uc}-X_{up})^2}{\sigma_u} \right] & \text{if } j, \ell > G, \\ -\sum_{u=1}^U \frac{X_{uj}X_{u\ell}}{\sigma_u} & \text{if } j, \ell \leq G, \\ -\frac{(V_{uc}-X_{up})}{\sigma_u^2} X_{uj}X_{u\ell} & \text{if } j \leq G, \ell > G, \\ -\frac{(V_{uc}-X_{up})}{\sigma_u^2} X_{uj}X_{u\ell} & \text{if } \ell \leq G, j > G. \end{cases} \quad \sigma = x_{u1}\lambda_1 + \dots + x_{uG}\lambda_G$$

The inverse of the expectation of this is simply

$$\begin{bmatrix} -(Z'Z)^{-1} & 0 \\ 0 & -2(W'W)^{-1} \end{bmatrix},$$

where $Z_u = \frac{X_u}{\sigma_u^{1/2}}$ and $W_u = \frac{X_u}{\sigma_u}$, for $\sigma_u > 0$.

This expectation is negative definite if the $X'X$ matrix is of full rank and if $\sigma_u > 0$ for all u , which suggests that the use of the Newton-Ralphson algorithm would probably give good results in maximizing the likelihood. The problem is that while the expectation of the inverse of the matrix of second partials is certainly negative definite at the true value of the parameters, neither the inverse of the matrix of second partials, nor its expectation, will in general be negative definite throughout the parameter space. There are several methods which are commonly used to handle this problem. One which we used in our programming was to modify the step size, so that the full Newton-Ralphson step was not taken when large changes in the coefficients were indicated by the algorithm (rather, a half-step or fourth-step was taken). We also used “good” starting estimates (generally the least square estimates) for our procedure.

²⁵Here the probabilities are the first G parameters and the variance parameters are the last G . We also use this matrix to derive the information matrix.

9 Comparing this model with an individual level data set

Our estimation was done with census tract level data for Los Angeles County, California for the June 1988 presidential primary election. ²⁶ We subdivided the latino and white populations by whether or not they lived in areas that were at least 50 % black. Running the Newton-Ralphson algorithm without this subdivision caused the model to fail to converge.

To compare our estimates with individual level estimates, we examined data from an exit poll conducted by the *Los Angeles Times* on the day of the 1988 Democratic presidential primary. There are 4,705 respondents from the state of California. We examined the subset of 2,193 respondents who said they voted in the Democratic primary. The ethnic breakdown of the candidates' support are provided in the regression results which follow.

Table 2 shows the results of the regression with dependent variable "Votes for Jackson".

²⁶Census tracts (with average population about 4000) are likely to be too large a unit of aggregation to expect to be able to satisfy the homogeneity conditions. They do allow the testing of our model in a racially polarized election.

Table 2 - SUPPORT FOR JACKSON

1988 California Primary	LA TIMES	LA County	Std. Errors (LA County)
COEFFICIENTS			
BLACK	.938	.943	.0037
ASIAN	.448	.590	.0160
LATINO	.433	.388	.1852
- living in black areas	n	.720	.0102
- living in non-black areas	n	.289	.0016
WHITE	.293	.189	.1836
- living in black areas	n	.257	.0042
- living in non-black areas	n	.178	.0005
VARIANCES			
BLACK	n	.77	.1482
ASIAN	n	11.84	19.9082
LATINO	n	1.30	.8571
- living in black areas	n	1.39	.8661
- living in non-black areas	n	1.28	.2730
WHITE	n	4.86	2.3704
- living in black areas	n	16.92	34.9162
- living in non-black areas	n	2.81	.0475
Overall Wald for model	n	97041.13	
Wald for location parameters	n	96273.68	
Wald for variance parameters	n	694.81	

n = not applicable or available for this type of data

Table 2 demonstrates several things about the model. First, the hypothesis of the estimated variances equalling the expected variances is rejected at any level of significance by the Wald test, as might be expected with comparing the estimation of the variance parameters from the model with the theoretical values, $p_g(1 - p_g)$. Of interest, however, is the fact that the estimated probabilities are very close to the exit poll (which was a statewide poll, whereas our estimates are for LA County only). One likely problem with our specification is that we utilized the ratio of the various ethnic and racial groups in the population for our creation of the partitions—thus our partitions are doubtlessly in error.²⁷

It is obvious a great deal of work needs to be done on how the model behaves when there is misspecification, because in some sense this model clearly is “good” when compared with the Los Angeles Times exit poll.

10 How to use this measure to assess racially polarized voting.

We conclude by proposing a new measure for the estimation of racially polarized voting, based on what has been developed here. Ideally, the measure we propose is one which:

1. satisfies the partition condition, (Section 1.)
2. is not statistically rejected when the homogeneity condition is tested, (Sections 1 and 7)
3. gives a coefficient for each racial or ethnic group by summing over all of the partitioned subdivisions of the electorate which that group includes, (Section 2)
4. and is used in conjunction with the conditions set out by Justice Brennan in *Thornberg v. Gingles* (1986)²⁸ or similar conditions which determine

²⁷Also, Los Angeles County, which has a large number of aliens in its heavily Latino areas, would by this method of creating partitions have too high a proportion of Latino voters. This is probably the reason that the model failed to converge when the Latinos and whites were not split by black areas.

²⁸106 S.Ct. 2767:

whether or not an electoral system complies with the amended Voting Rights Act.²⁹

Given a correct measure, the Congress may want to consider mandating a procedure for vote dilution cases which in some way recognizes the importance of the probability estimates by the various groups for the various candidates (i.e., $p_{E_1c} - p_{E_2c}$, as derived in Section 2, where E_1 and E_2 are different ethnic or racial groups and p_{gc} is the voting tendency of the specified group.)

11 Conclusion

Where previous Voting Rights cases have tended to congregate in the bi-racial South, the most recent cases involve multi-racial metropolitan areas in the northeast and southwest parts of the country. This movement will undoubtedly be the cause of much disagreement in the correct way to measure voting polarization. We have proposed a method which should address most of the controversies head on. While we have pointed out problems with estimation methods previously used and have proposed the homogeneity and partition conditions as necessary conditions for correct estimation, our conclusion is not that all previous estimation is flawed. Consider the voting rights cases of the South, where the electoral behavior of blacks and whites in races where some of the candidates were black was given special attention. In those situations, it is more than likely that the necessary conditions were satisfied.

The real concern here is for present and future estimation where the identification of group voting behavior is not so clear. In contemporary America, it cannot be taken for granted that all members of a protected minority group

- The minority group must be able to demonstrate that it is sufficiently large and geographically compact to constitute a majority in a single member district.
- The minority group must show that it is politically cohesive.
- The minority must be able to demonstrate that the white majority votes sufficiently as a bloc to enable it usually to defeat the minority preferred candidate.

²⁹This concept has been worked with most extensively in Grofman, Migalski and Noviello (1985).

have the same electoral preferences. Estimates of group voting behavior based on such an assumption, ignoring the homogeneity condition, will be of questionable validity. These estimates will ignore the fact that voting decisions are made by individuals, and individual behavior cannot necessarily be modelled through ethnic variables alone. We have offered a method which allows a determination of what categorizations are sufficient.

We have shown that this measure is one which is based upon an intuitively plausible assumption about individual behavior, with probabilistic reasoning being applied to derive conditions which aggregate estimates must satisfy. We have also seen, in our empirical example, that our method of estimating provides fairly accurate results even when the conditions we have set forth are violated. Further research is obviously needed to understand exactly what conditions this can happen under. In any case, we remain confident that the theoretical underpinnings of our method are sound and that any other model which purports to measure racial polarization in an electorate must also be based upon the behavior of individuals in that electorate. In particular, the correlation coefficient has clearly been shown to be deficient and should no longer be used. In addition, regression methods should not be used without examination under our probabilistic framework.

We quite honestly believe that further work needs to be done on this model before we would recommend it as being accepted as the one and only method of determining racially polarized voting. (One possible extension, suggested by Gary King, would be to allow the probabilities to be randomly distributed, perhaps through the use of a beta distribution.) We do believe that all future analyses utilizing regression should be estimated with this model, as well as with regression. And it is clear, also, that a probabilistic framework of the type we have set up, with the emphasis on deriving aggregate methods of analysis from assumptions on the behavior of individuals, is in the long run the only type of framework that is defensible, either in a courtroom or in scholarly research. The establishment of racially polarized voting must, in the final analyses, be based on the voting behavior of individuals of that minority.

APPENDIX ELECTORATES IN GREATER DETAIL

SAMPLE ELECTORATE	Electorate U	Electorate V
Covariances		
$E_1 I_1$, E_1	-4.25	18.75
$E_1 I_1$, E_2	4.25	-18.75
$E_1 I_1$, I_2	-4.50	-77.00
$E_1 I_2$, E_1	10.75	-12.25
$E_1 I_2$, E_2	-10.75	12.25
$E_1 I_2$, I_2	12.25	51.00
$E_2 I_1$, E_1	-3.50	7.25
$E_2 I_1$, E_2	3.50	-7.25
$E_2 I_1$, I_2	-5.75	-28.69
$E_2 I_2$, E_1	-3.00	-13.75
$E_2 I_2$, E_2	3.00	13.75
$E_2 I_2$, I_2	-2.00	54.69
Variances		
E_1	6.50	6.5
E_2	6.50	6.5
I_2	10.25	105.7

A Derivation of the correlation coefficient when homogeneity condition is satisfied.

We assume that $X_1 + X_2 = 1$, $E(\epsilon_1) = 0$, $E(\epsilon_2) = 0$ and that ϵ_1 and ϵ_2 are independent of each other, X_1 and X_2 .

By the definition of the correlation coefficient,

$$\text{Corr}(X_1, V_1) = \frac{\text{Cov}(X_1, V_1)}{[\text{Var}(X_1)\text{Var}(V_1)]^{1/2}}.$$

The vote for candidate 1 is defined as:

$$\begin{aligned} V_1 &= X_1(p_1 + \epsilon_1) + X_2(p_2 + \epsilon_2) \\ &= X_1p_1 + X_2p_2 + X_1\epsilon_1 + X_2\epsilon_2. \end{aligned}$$

Which gives,

$$\text{Var}(V_1) = \text{Var}(X_1p_1 + X_2p_2) + \text{Var}(X_1\epsilon_1 + X_2\epsilon_2),$$

since $\text{Cov}(X_1p_1 + X_2p_2, X_1\epsilon_1 + X_2\epsilon_2) = 0$, and,

$$\text{Var}(X_1\epsilon_1 + X_2\epsilon_2) = \text{Var}(X_1\epsilon_1) + \text{Var}(X_2\epsilon_2),$$

since we assume $\text{Cov}(X_1\epsilon_1, X_2\epsilon_2) = 0$. From the definition of variance, we have,

$$\begin{aligned} \text{Var}(X_1\epsilon_1) &= E(X_1^2)E(\epsilon_1^2) - [E(X_1)]^2[E(\epsilon_1)]^2, \\ &= E(X_1^2)E(\epsilon_1^2), \end{aligned}$$

$$\text{Var}(X_2\epsilon_2) = E(X_2^2)E(\epsilon_2^2),$$

and

$$\text{Var}(X_1p_1 + X_2p_2) = p_1^2\text{Var}(X_1) + p_2^2\text{Var}(X_2) + 2p_1p_2\text{Cov}(X_1, X_2).$$

We also know that

$$\text{Var}(X_1) = \text{Var}(1 - X_2) = \text{Var}(-X_2) = \text{Var}(X_2),$$

and

$$\text{Cov}(X_1, X_2) = \text{Cov}(X_1, (1 - X_1)) = -\text{Var}(X_1).$$

Now,

$$\begin{aligned} \text{Cov}(X_1, V_1) &= \text{Cov}(X_1, X_1p_1 + X_2p_2 + X_1\epsilon_1 + X_2\epsilon_2), \\ &= \text{Cov}(X_1, X_1p_1 + X_2p_2), \\ &\quad (\text{since } \text{Cov}(X_1, X_1\epsilon_1 + X_2\epsilon_2) = 0) \\ &= \text{Cov}(X_1, X_1p_1 + (1 - X_1)p_2), \\ &= \text{Cov}(X_1, X_1(p_1 - p_2) + p_2), \\ &= (p_1 - p_2)\text{Cov}(X_1, X_1), \\ &= (p_1 - p_2)\text{Var}(X_1). \end{aligned}$$

We now have all of the component parts of the correlation coefficient in terms of p_i , X_i and ϵ_i , which we substitute into the definition,

$$\begin{aligned}\text{Corr}(X_1, V_1) &= \frac{(p_1 - p_2)\text{Var}(X_1)}{[\text{Var}(X_1)]^{1/2} [(p_1 - p_2)^2 \text{Var}(X_1) + E(X_1^2)E(\epsilon_1^2) + E(X_2^2)E(\epsilon_2^2)]^{1/2}} \\ \text{Corr}(X_1, V_1) &= \frac{(p_1 - p_2)}{\left[(p_1 - p_2)^2 + \frac{E(X_1^2)}{\text{Var}(X_1)} E(\epsilon_1^2) + \frac{E(X_2^2)}{\text{Var}(X_1)} E(\epsilon_1^2) \right]^{1/2}}\end{aligned}$$

If our assumption of homogeneity is satisfied, from the derivation in section 3, we have

$$\begin{aligned}E(\epsilon_1^2) &= p_1(1 - p_1), \text{ and} \\ E(\epsilon_2^2) &= p_2(1 - p_2).\end{aligned}$$

$$\text{Let } \frac{E(X_1^2)}{\text{Var}(X_1)} = a_1 \text{ and } \frac{E(X_2^2)}{\text{Var}(X_2)} = a_2.$$

Then

$$\text{Corr}(X_1, V_1) = \frac{(p_1 - p_2)}{[(p_1 - p_2)^2 + a_1 p_1(1 - p_1) + a_2 p_2(1 - p_2)]^{1/2}}.$$

Note that this derivation depends completely upon assumptions of our model being satisfied—otherwise, the covariance terms are not zero and they must be included in the expression. In addition, notice that if our assumptions are satisfied:

$$\frac{\text{Cov}(X_1, V_1)}{\text{Var}(X_1)} = p_{1c} - p_{2c}.$$

B An example of the range of possible values that the correlation coefficient can produce.

Let $p_2 = 0$, then,

$$\text{Corr}(X_1, V_1) = \frac{p_1}{[p_1^2 + a_1 p_1(1 - p_1)]^{1/2}}.$$

Dividing through by p_1 gives us:

$$\text{Corr}(X_1, V_1) = \frac{1}{\left[1 + a_1 \frac{(1-p_1)}{p_1}\right]^{1/2}}.$$

Let $p_1 = .5$, then,

$$\text{Corr}(X_1, V_1) = \frac{1}{(1 + a_1)^{1/2}}.$$

$$a_1 = \frac{E(X_1^2)}{\text{Var}(X_1^2)} = \frac{E(X_1^2)}{E(X_1^2) - [E(X_1)]^2} = \frac{1}{1 - \frac{[E(X_1)]^2}{E(X_1^2)}}$$

To estimate the possible range of the correlation coefficient we need to know some bounds on a_1 . X_1 is a proportion (as defined in Section 5) so we know X_1 is between zero and one. Thus, we have,

$$E(X_1^2) \leq E(X_1) \leq 1 \text{ and } E(X_1^2) \geq [E(X_1)]^2,$$

where the last relation is by the Cauchy-Schwartz inequality. Then,

$$E(X_1^2) \leq E(X_1) \leq [E(X_1^2)]^{1/2}.$$

Let $E(X_1^2) = .25$. Then the above implies,

$$.25 \leq E(X_1) \leq .5,$$

which implies,

$$.0625 \leq [E(X_1)]^2 \leq .25$$

and

$$1 \geq \frac{[E(X_1)]^2}{E(X_1^2)} \geq \frac{1}{4}.$$

Plugging in to the equation for the correlation coefficient derived above we get:

$$0 \leq \text{Corr}(X_1, V_1) \leq .65.$$

Similarly, letting $E(X_1^2) = .64$ implies $0 \leq \text{Corr}(X_1, V_1) \leq .625$.

In this case we show that the correlation coefficient can be as high as .65, which some consider to be evidence of significant polarization, when a majority of neither group supports the candidate in question. It should be evident that a range of values of p_1, p_2 and the moments of X will produce similar troubling results.

C Estimation Program

With the statistical program, SPSS-X (Version 3.0 is what we used) we can estimate the varying parameters model using the constrained non-linear regression function. Below is a version of the SPSS-X command file needed for estimation. The “comment” statements are included to help the reader keep track of our variable names and what our intention is at each stage. Our space restrictions will lead the interested reader to examine the SPSS-X manual for further information.

```
file handle insys/name='rbv_p88.sys'
comment
    This is the name of the input file.
get file=insys/
comment
    The unit of observation is the voting precinct or a census tract.
    Since we are looking at a (Democratic presidential, Los Angeles
    County, 1988) primary, we break this electorate down into the
    relevant ethnic and racial categories using election returns and
    census data.

(nblack = number of blacks)
(demvot = number of voters participating in Democratic primary)
(totvot = number of voters on primary day)

compute black    = demvot/totvot * nblack
compute latino   = demvot/totvot * nlatino
compute amerind  = demvot/totvot * namerind
compute asian    = demvot/totvot * nasian
compute other    = demvot/totvot * nother
compute white    = demvot/totvot * nwhite

comment
    When using the constrained non-linear regression method, initial
```

estimates of the support of the racial groups for a particular candidate are given with the model program command. This gives SPSS-X a starting point from where it can converge to the true values. From the binomial theorem, we can also give corresponding initial values for the variance of each parameter.

```
(cblack = initial estimate of black support for Dukakis)
(sblack = initial estimate variance of black support for Dukakis)
(prsdm1$ = votes for Dukakis)
```

model program

```
cblack= .01 clatino= .50 camerind= .50 casian= .50 cother= .50
cwhite= .70 sblack= .099 slatino= .25 samerind= .25 sasian= .25
sother= .25 swhite= .21
```

comment

"Pred" gives the predicted value of the dependent variable (In our example, that is "Vote for Dukakis" or "Vote for Jackson"). "Loss" is equivalent to a loss function in maximum likelihood estimation, which is a representation of the difference between the predicted and actual values of the dependent variable. We get our estimates of racially polarized voting through SPSS-X's successive attempts to minimize this loss function.

compute pred =

```
(black * cblack) + (latino * clatino) + (amerind * camerind)
+ (asian * casian) + (other * cother) + (white * cwhite)
```

compute loss=

```
ln ((black * sblack) + (latino * slatino) + (amerind * samerind)
+ (asian * sasian) + (other * sother) + (white * swhite))
+ ((prsdm1$ - pred)* (prsdm1$ - pred))/
(black * sblack) + (latino * slatino) + (amerind * samerind)
+ (asian * sasian) + (other * sother) + (white * swhite))
```

comment

"Cnlr" is the constrained non-linear regression command. "prsdm1\$" (votes for Dukakis) is the dependent variable. Variables listed after "with" are the independent variables. There is no constant term as the groups listed are collectively exhaustive of the electorate. The "bounds" subcommand, gives the necessary parameter limits.

```
cnlr prsdm1$ with black latino amerind asian other white
```

```
  /bounds
```

```
    .0001 < cblack    < .9999    ;
```

```
    .0001 < clatino   < .9999    ;
```

```
    .0001 < camerind  < .9999    ;
```

```
    .0001 < casian    < .9999    ;
```

```
    .0001 < cother    < .9999    ;
```

```
    .0001 < cwhite    < .9999    ;
```

```
    .0001 < sblack    ;
```

```
    .0001 < slatino   ;
```

```
    .0001 < samerind  ;
```

```
    .0001 < sasian    ;
```

```
    .0001 < sother    ;
```

```
    .0001 < swhite    ;
```

```
  /loss=loss
```

```
  /criteria steplimit .01
```

finish

D Bibliography.

1. Bullock, Charles S. III and Susan A. MacManus. "Measuring Racial Bloc Voting is Difficult for Small Jurisdictions". *National Civic Review*, July/ August, 1984
2. Chow, Gregory C.. "Random and Changing Coefficient Models" in *Handbook of Econometrics: Volume 2*. (Amsterdam, 1984).
3. Engstrom, Richard L. and Michael D. MacDonald. "Quantitative Evidence in Vote Dilution Litigation, Part II: Minority Coalitions and Multivariate Analyses", *Urban Lawyer*, Winter 1987, Vol. 19 # 1.
4. Feller, William. *An Introduction to Probability Theory and its Applications: Volume I*. (New York, 1950).
5. Grofman, Bernard, Michael Migalski and Nicholas Noivello. "The Totality of Circumstances" test in Section 2 of the 1982 Extention of the Voting Rights Act: A Social Science Perspective", *Law and Policy*, Vol. 1, Number 2 (1985).
6. Grofman, Bernard, Michael Migalski and Nicholas Noviello. "Effects of Multimember Districts on Black Representation in State Legislatures", *The Review of Black Political Economy*, 1986, Vol. 14, Number 4.
7. Swamy, P.A.V.B.. *Statistical Inference in Random Coefficient Regression Models*. (New York, 1971).
8. Thornberg v. Gingles (1986), 106 S.Ct. 2767.
9. United States Senate. Senate Report # 97-417, 97th Congress, Second Session 28 (1982). U.S. Code Cong. and Admin. News 1982.